Leak Localization in Drinking Water Distribution Networks using Structured Residuals

Albert Rosich[1,*], Vicenç Puig[2] and Myrna V. Casillas[3]

[1] *Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg,* e-mail: albert.rosich@uni.lu
[2] *Advanced Control Systems (SAC), Universitat Politècnica de Catalunya (UPC), Pau Gargallo 5, 08028 Barcelona, Spain,* e-mail: vicenc.puig@upc.edu
[3] *Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, 64849 México,* e-mail: mv.casillas.phd.mty@itesm.mx

SUMMARY

In this paper, a new model based approach to leakage localization in drinking water distribution networks is proposed based on generating a set of structured residuals. The proposed method is suitable for water distribution networks because the model is given as a system of non-linear equations with no explicit solution that makes impossible to derive analytical expressions for the residuals. For this reason, the residual evaluation is based on a numerical method based on an enhanced Newton-Raphson algorithm. Moreover, because this type of networks presents a high degree of interconnection between nodes, the primary residuals, obtained directly from the model and measurements, do not allow to easily isolate the faults. To address this problem, the proposed approach allows to derive new structured sets of residuals so that leaks are decoupled, which improves the localization of leaks with respect to primary residuals. Finally, the proposed approach is tested on a real water network model as well as compared and integrated with a previous existing approach based on directional residuals.

Received . . .

1. INTRODUCTION

Water loss in drinking wate distribution networks, caused by accidentally pipe bursts or intentionally breakdowns, is an issue of great concern for water utilities, strongly linked with operational costs and water resources savings. Continuous improvements in water loss management are being applied and new technologies are developed to achieve higher levels of efficiency. Usually, a leakage detection method in a DMA (District Metered Area) starts by analyzing input flow data, such as minimum night flows and consumer metering data [1, 2]. Once a DMA is identified to have a leakage, various techniques are used to locate it for pipe replacement or repair. Methods for locating leaks range from ground-penetrating radar to acoustic listening devices or physical inspection [3] [4]. Some of these techniques require isolating and shutting down part of the system. The whole process could take weeks or months with a significant volume of water wasted.

Recently, techniques based on standard theory of model-based diagnosis have been proposed in order to detect and locate leaks from pressure measurement [5], allowing more effective and less costly search in situ. These techniques exploit the knowledge provided by the model in order to

*Correspondence to: Interdisciplinary Centre for Security, Reliability and Trust in the University of Luxembourg, Luxembourg, e-mail: {albert.rosich}@uni.lu

†Please ensure that you use the most up to date class file, available from the ACS Home Page at

monitor the behaviour of the system by means of residual computation and evaluation. The residual expressions are typically obtained from analytical redundancy relations involved in the system model and its input variables are the system observations (e.g. on-line control signals and/or sensor measurements). To detect faults, residuals are compared with a threshold value (zero in ideal case or almost zero in real case). Fault localization can be performed when residuals behave differently under different faults.

In the literature, two different approaches to construct residual sets with the desired isolability properties have been proposed. One approach is based on designing a vector of *structured residuals* [6]. Each residual is designed to be sensitive to a subset of faults, whilst remaining insensitive to the remaining faults. The design procedure consists of two steps: the first step is to specify the sensitivity and insensitivity relationships between residuals and faults according to the assigned isolation task, while the second is to design a set of residual generators according the previous relationships. The fault isolation problem then consists in a separately threshold test for each residual using a decision table. An alternative way of achieving the isolability of faults is to design a vector of *directional residuals* [6], which lies in a fixed and fault-specified direction in the residual space, in response to a particular fault. The fault isolation problem consists in determining which of the known fault directions, called *fault signatures*, the generated residual vector lies the closest to.

The water network model is described by a set of non-linear equations with non-explicit solution, therefore standard techniques of model-based diagnosis can not be straightforwardly applied. Directional residuals methods have already been suggested for water distribution networks [5, 7] where the directions are determined by means of simulations. However, the obtained directions strongly depend on the operation point and the leak magnitude chosen in the simulations, which causes an increase of the false leak location ratio as the network conditions differ from the simulation ones.

Alternatively, the aim of this paper is to develop an efficient method based on structured residuals in order to detect and locate leaks in a water distribution network. This new proposed method differs from previous one cited above in the sense that the class of residuals designed in this paper presents suitable structural properties for leak localization. Moreover, an algorithm that numerically computes the residuals is presented to alleviate the problem due to the fact that the obtained residuals are given as a system of non-linear equations with no explicit solution making impossible to derive analytical expressions for the residuals.

The proposed approach will be compared with the directional residual approach reported in [7] in a real water network proposed as a case study. As a result of this comparison, the complementarity of two approaches will be highlighted and an integrated approach will be suggested such that the leak location ratio will be significantly enhanced because of the properties from both classes are used.

The organization of the paper is as follows. In Section 2, the water distribution network model for the no leaky case as well as for the leaky case is presented. In addition, a numerical method to simulate the water model is briefly reviewed. In Section 3, the paper is motivated by showing the drawbacks of the existing model-based leak localization methods based on directional residuals. In Section 4, the principles of the new proposed leak localization methodology are introduced, and the algorithm to compute the structured residuals is introduced as well. In Section 5, a diagnosis analysis is presented in order to know beforehand which leaks can be detected and located. The case study and results obtained through the proposed approach are shown in Section 6. Finally, Section 7 concludes the paper and suggests some future research lines to extend the proposed methodology.

2. WATER NETWORK MODEL

The water network model is typically defined by a set of N nodes representing the junctions or reservoirs, and a set of edges E representing the pipes. Let q_e for each $e \in E$ with $e = (i, j)$ be the flows from node i to node j, and let h_n for each $n \in N$ be the pressures. The set of pressures is partitioned into two disjoint subsets: the subset of junction pressures $U \subset N$ and the subset of reservoir pressures $K \subset N$ (pressures at reservoirs are fixed and known by definition).

The network in steady state fulfills the following stationary point conditions:

- Flow balance conservation,

$$\sum_{n,j} q_{nj} = d_n \qquad \text{for all } n \in U \tag{1}$$

where d_n is the demand at node n.

- Hansen-Williams pressure loss functions,

$$h_i - h_j = r_e q_e |q_e|^{\gamma_e} \qquad \text{for all } e \in E; e = (i,j) \tag{2}$$

where r_e is a pipe parameter which depends on diameter, roughness and length, and γ_e is the flow exponent parameter.

- Reservoir pressure condition,

$$h_n = h_n^* \qquad \text{for all } n \in K \tag{3}$$

where h_n^* is a known pressure value.

Let p_u, p_k and p_e be the number of junction nodes, reservoirs and flows, respectively (i.e. $p_u = |U|$, $p_k = |K|$ and $p_e = |E|$), and also let $\mathbf{h} = (h_1, \ldots, h_{p_u})^T$, $\mathbf{h_0} = (h_1^*, \ldots, h_{p_k}^*)^T$, $\mathbf{q} = (q_1, \ldots, q_{p_e})^T$ and $\mathbf{d} = (d_1, \ldots, d_{p_u})^T$ denote the vector of junction node pressures, reservoir pressures, flows and known demands, respectively. Then, the water network model can be formulated in matrix form as follows:

$$\begin{pmatrix} A_{11}(\mathbf{q}) & A_{12} \\ A_{21} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{h} \end{pmatrix} = \begin{pmatrix} -A_{10}\mathbf{h_0} \\ \mathbf{d} \end{pmatrix} \tag{4}$$

where $A_{11}(\mathbf{q}) = diag(r_1|q_1|^{\gamma_1}, \ldots, r_{p_e}|q_{p_e}|^{\gamma_{p_e}})$, $A_{12} = A_{21}^T$, and $A_{10} = A_{01}^T$. The matrices A_{21} and A_{01} are the incidence matrices obtained from the network graph when only junction and reservoir nodes are respectively considered. The resulting water network model has $\mathbf{q}$ and $\mathbf{h}$ as unknowns, p_e non-linear equations (pressure losses) and p_u linear equations (flow balances).

2.1. *Adding leaks to the model*

A water leak could theoretically appear at any point of the network pipes which would imply that the graph structure of the model should be modified by adding a new junction node at the leaking point as well as new parameters should be derived for the leaking pipe. For this reason, the modeling of any possible leak becomes difficult and unfeasible in practice. To mitigate this problem, it is usually assumed that leaks can only appear at existing nodes [8]. Therefore, the number of leaks to be considered can be constrained by the number of existing junction nodes in the network. In addition, no new parameter estimation needs to be done since no change in the pipe is assumed. In this section, the model in (4) is extended by taking into account the assumed possible leaks.

Let $\mathbf{l} = (l_1, \ldots, l_{p_u})^T$ be the vector of possible leaks in the water network, then the model (4) can be extended to also include leaks in nodes as

$$\begin{pmatrix} A_{11}(\mathbf{q}) & A_{12} & 0 \\ A_{21} & 0 & -I \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{h} \\ \mathbf{l} \end{pmatrix} = \begin{pmatrix} -A_{10}\mathbf{h_0} \\ \mathbf{d} \end{pmatrix} \tag{5}$$

Usually, this model is implicitly used to study leak sensitivity analysis by means of simulations (for more details see [5, 7]). For instance, assume that the effect of leak i is studied for a specific known value l_i. Then, according to (5), this can be done by solving model (4) with $\mathbf{d} = (d_1, \ldots, d_i + l_i, \ldots, d_{p_u})^T$.

DOI: 10.1002/acs

2.2. Model Solution

The model in (4) is not linear and cannot be solved analytically. Therefore numerical tools are needed in order to provide a solution in terms of flows, $\mathbf{q}$, and pressures $\mathbf{h}$. In [9], it is proved that the solution for a water network exists and is unique. This proof relies on the following non-linear optimization problem,

$$\min\Big\{\sum_{i=1}^{n_e}\Big(\int_0^{q_i} r_i|t|^{\gamma_i}dt\Big)+\sum_{j=1}^{n_k}h_j^*\sum_{i=1}^{n_e}A_{10_{(i,j)}}q_i\Big\} \tag{6}$$

subject to:

$$\sum_{i=1}^{n_e}A_{21_{(j,i)}}q_i-d_j=0 \qquad \text{for } j=\{1,\dots,p_u\} \tag{7}$$

that is known as *Content Model* [10], and its solution is given by (4). Since all r_i and γ_i are positives, the minimization problem is convex which means that the solution of both, the *Content Model* and the water network, exists and is unique.

Besides, [9] presents a numerical method based on the recursive Newton-Raphson algorithm to solve the non-linear water model (4). The algorithm can be summarized with the following two recursive equations:

$$\mathbf{h}_{(k+1)}=-\left(A_{21}N^{-1}A_{11}(\mathbf{q})^{-1}A_{12}\right)^{-1}\Big(A_{21}N^{-1}(\mathbf{q}_{(k)}+A_{11}(\mathbf{q})^{-1}A_{10}\mathbf{h_0})+(\mathbf{d}-A_{21}\mathbf{q_{(k)}})\Big) \tag{8}$$

$$\mathbf{q}_{(k+1)}=(I-N^{-1})\mathbf{q}_{(k)}-N^{-1}A_{11}(\mathbf{q})(A_{12}\mathbf{h}_{(k)})+A_{10}\mathbf{h_0}) \tag{9}$$

where $N=diag(\gamma_1,\gamma_2,\dots,\gamma_{p_e})$ and $A_{11}(\mathbf{q})$ is computed using $\mathbf{q}_{(k)}$. Here, it will be assumed that no flow $q\in\mathbf{q}_{(k)}$ is zero (null flows would lead to an ill conditioned problem where network reconfiguration techniques should be used). The algorithm is initialized with any guess values, $\mathbf{h}_{(0)}$ and $\mathbf{q}_{(0)}$, and the solution is found by recursively applying equations (8)-(9) until a preestablished degree of tolerance is achieved.

This method is used in the EPANET simulator [11] where large water networks can be efficiently solved.

3. MOTIVATION

The most recent model-based methods used to locate leaks in water distribution networks analyze the effect that possible leaks have in the pressure measurements obtained from the available sensors installed by means of a leak sensitivity analysis (see for example, [8, 5]). There exist several approaches to derive directional residuals from the sensitivity matrix S, see [7]. According to [7], the most efficient method is the *Angle between Vectors* (or *Projection*) method which is next reviewed describing the weak points that have motivated the approach proposed in this paper.

Let $M\subseteq K$ be the set of measured junction pressures and p_m the number of pressure sensors installed in the network (i.e. $p_m=|M|$), then p_m primary residuals can be computed as follows

$$r_i=m_i-\hat{h}_i \qquad \text{for } i\in M \tag{10}$$

where m_i is the measured pressure value in node i, and $\hat{h}_i$ is the corresponding model-based estimation computed, for example, using EPANET. Note that the residuals should be zero (or nearly zero) when there is no leak while they should be different from zero when a leak is present, i.e. leak detection is achieved by simply evaluating how far residuals are from zero. Specifically, a pressure sensitivity vector is computed for each possible leak as follows:

$$\mathbf{s}_{l_j}=\frac{\hat{\mathbf{h}}_{l_j}-\hat{\mathbf{h}}}{l_j} \qquad \text{for } j=\{1,\dots,n_l\} \tag{11}$$

where $\hat{\mathbf{h}} = (\hat{h}_1, \ldots, h_{p_m})^T$ is computed by assuming no leak (model (4) is used), while $\hat{\mathbf{h}}_{l_j}$ is a similar vector computed by assuming a nominal leak l_j (model (5) is used with $\mathbf{l} = (0, \ldots, l_j, \ldots, 0)^T$). Therefore, the i-element of the sensitivity vector $\mathbf{s}_{l_j}$ stands for the expected deviation of the preassure measurement h_i under the nominal leak l_j effect.

It should be noted that an analytical computation of vector $\mathbf{s}_{l_j}$ is not feasible since, as shown in equations (8)-(9), the pressure deviations depends non-linearly on the flows which are not measured neither can be computed in the case that leaks are present in the real water network.

The leak localization method is based on evaluating the projection of the residual vector $\mathbf{r} = (r_1, \ldots, r_{p_m})^T$ on the sensitivity vector $\mathbf{s}_{l_j}$ for each possible leak, i.e.

$$\alpha_{l_j} = \frac{\mathbf{r}^T \mathbf{s}_{l_j}}{\|\mathbf{r}\| \|\mathbf{s}_{l_j}\|} \qquad \text{for } j = \{1, \ldots, n_l\} \tag{12}$$

Notice that α_{l_j} is equal to cosine of the angle between the to vectors $\mathbf{r}$ and $\mathbf{s_{l_j}}$. This is why this leak localisation method is named as the *Angle between Vectors* method in [7].

Since vectors are normalized in (12), the scalar value α_{l_j} must be in the interval $[-1\ 1]$, where $\alpha_{l_j} = 1$ means that the residual direction is the same as $\mathbf{s}_{l_j}$ while $\alpha_{l_j} = -1$ indicates opposite direction. Note that the values of variables α_{l_j} can be regarded as a plausibility degree for leak l_j being present in the water network. Therefore, the leak with α value closer to one is chosen by the diagnosis system as the candidate leak.

It is worth to point out that in this method no measured pressure is used to simulate the model (only reservoir pressures, h^*, are used) whereas measured pressures are used for generating the residual (10), thus the number of residuals is the same as the number of available sensors. However, as mentioned before, the residual sensitivity is not constant but depends on the leak magnitude, which makes this method less effective in case that different nominal leak magnitudes from the ones chosen in (11) are present in the water network.

To mitigate this problem, an alternative approach is presented in the next section. By using this alternative method, each residual becomes decoupled from certain desired leaks and thus the localization task does not depend on the leak magnitude.

4. STRUCTURED RESIDUALS IMPLEMENTATION FOR LEAK LOCALIZATION

The residual generator proposed in this new paper has the same form as residual (10), i.e.

$$r = m_k - \tilde{h}_k \qquad \text{for } k \in M \tag{13}$$

where $\tilde{h}_k$ is also a model-based estimation of the pressure. However, the procedure to obtain $\tilde{h}_k$ is different from the one used to obtain $\hat{h}_i$ in (10).

4.1. *Leak decoupling*

As it was shown in Section 2.2, the system of equations describing the water network model (4) can be solved and the solution is unique. Now, consider that the pressure in junction node $i \in U$ is measured, i.e., $i \in M$. This means that $h_i \in \mathbf{h}$ becomes known and thus it can be removed from the vector of unknown pressures. Then, $\tilde{\mathbf{h}}[i] = (h_1, \ldots, h_{i-1}, h_{i+1}, \ldots, h_{p_u})^T$ denotes the new vector of unknown pressures, while the corresponding measured value m_i can be added to the vector of known pressures. Thus, $\tilde{\mathbf{h}}_\mathbf{0}[i] = (h_1^*, \ldots, h_{p_k}^*, m_i)^T$ denotes the new vector of known pressures values. Proceeding in this way, the solution space is reduced in one dimension which means that now one equation is redundant and therefore not needed to compute the solution.

Proposition 1
Given the water network model in (4) with no null flows and assuming that one junction node pressure in measured, then the model system equations can be solved (8)-(9) if any row j from matrix A_{21} and its corresponding demand d_j are removed.

Proof
First, note that in order to keep the notation consistent in (8)-(9) when the pressure of the i-th node is known, the i-th columns of matrix A_{12} must be removed and inserted into A_{10}. This only affects to the term $(A_{21}N^{-1}A_{11}(\mathbf{q})^{-1}A_{12})$ in (8) which needs to be invertible.

Matrices N and $A_{11}(\mathbf{q})$ are invertible (under the non-null flow assumption) and are not affected by the pressure measurement. Then, from graph theory, it is easy to show that matrix A_{21} is an $m \times n$ dimension matrix with $m \leq n$ and rank$(A_{21}) = m$. Consequently, matrix A_{12} is an $n \times m$ matrix with full column rank. This means that no rank deficient matrix is obtained by removing rows in A_{21} or columns in A_{12}. Therefore, the term will remain invertible as long as one row is removed from A_{21}.

Notice that the demand d_j must be also removed for coherence with the resulting model. □

The resulting solvable model when pressure in node $i \in U$ is measured can be represented as

$$\begin{pmatrix} A_{11}(\mathbf{q}) & \tilde{A}_{12}[i] \\ \tilde{A}_{21}[j] & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \tilde{\mathbf{h}}[i] \end{pmatrix} = \begin{pmatrix} -\tilde{A}_{10}[i]\tilde{\mathbf{h}}_{\mathbf{0}}[i] \\ \tilde{\mathbf{d}}[j] \end{pmatrix} \tag{14}$$

where $\tilde{A}_{12}[i]$ and $\tilde{A}_{21}[j]$ are the matrix obtained by removing the i-th column in A_{12} and the j-th row in A_{21}, respectively. $\tilde{A}_{10}[i]$ corresponds to the concatenation of matrix A_{10} with the i-th column of A_{12}, and $\tilde{\mathbf{d}}[j] = (d_1, \ldots, d_{j-1}, d_{j+1}, \ldots, d_{p_k})^T$.

It is important to note that the model obtained in (14) has no physical meaning and thus cannot be implemented in a water network simulator as EPANET. However, this new model presents an advantageous property compared with the former model.

Proposition 2
Given $\tilde{\mathbf{h}}[i]$ computed with model (14), then any residual $r = m_i - \tilde{h}_i$ for $\tilde{h}_i \in \tilde{\mathbf{h}}[i]$ is decoupled from leak l_j.

Proof
The model (14) is extended to also include the leak l_j, i.e.

$$\begin{pmatrix} A_{11}(\mathbf{q}) & \tilde{A}_{12}[i] & 0 \\ \tilde{A}_{21}[j] & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \tilde{\mathbf{h}}[i] \\ l_j \end{pmatrix} = \begin{pmatrix} -\tilde{A}_{10}[i]\tilde{\mathbf{h}}_{\mathbf{0}}[i] \\ \tilde{\mathbf{d}}[j] \end{pmatrix} \tag{15}$$

where the effect of the leak l_j is null because flow balance equation concerning node j has been removed according to the Preposition 1. Then, the proof consists in showing that the solution of the water model when there is a leak is also the solution of the model in (15).

Assume that $\bar{\mathbf{h}} = (\bar{h}_1, \ldots, \bar{h}_{p_u})^T$ and $\bar{\mathbf{q}} = (\bar{q}_1, \ldots, \bar{q}_{p_e})^T$ are the solutions of model (5) for any leak $\mathbf{l}$ such that $\mathbf{l} = (0, \ldots, 0, l_j, 0, \ldots, 0)^T$. Then, from (15), we have that

$$A_{11}(\bar{\mathbf{q}})\bar{\mathbf{q}} + \tilde{A}_{12}[i] \begin{pmatrix} \bar{h}_1 \\ \vdots \\ \bar{h}_{i-1} \\ \bar{h}_{i+1} \\ \vdots \\ \bar{h}_{p_u} \end{pmatrix} = -\tilde{A}_{10}[i] \begin{pmatrix} \tilde{\mathbf{h}}_{\mathbf{0}}[i] \\ \bar{h}_i \end{pmatrix} \tag{16}$$

must hold since the equations are the same as in (5), and

$$\tilde{A}_{21}[j]\bar{\mathbf{q}} = \tilde{\mathbf{d}}[j] \tag{17}$$

must hold since it does not depend on l_j. □

From this result, we can not guarantee that the residual sensitivity to leaks will be large enough. However, we are now able to implement a class of residuals which are insensitive to certain leaks, independently of its magnitude. In case of measuring more than one junction node pressure, the same result could be still applied by removing more than one column and row. Then, the residual would be decoupled from more than one leak. In this approach, sensors can be classified in two sets according its purpose: sensors for leak decoupling, and sensors for residual generation.

4.2. *Automatic residual computation*

Another advantage of this new proposed method for residual generation is that we can generate many residuals with few sensors. To illustrate this, consider now that there are two pressure sensors installed in the network. One sensor, placed in node $i \in U$, is used to decouple the leak from the residual according to (14), whereas another sensor, placed in node $k \in U$, is used to generate the residual according to (13) (for index notational consistency, we need to assume without loss of generality that $i > k$). Since the j-th row in (14) has been arbitrary chosen, we can use the same sensor to decouple all the leaks l_j for $j = \{1, \ldots, p_u\}$, obtaining p_u different pressure estimates, $\{\tilde{\mathbf{h}}[i]_1, \ldots, \tilde{\mathbf{h}}[i]_{p_u}\}$. Then, p_u residuals, each one insensitive to a different leak, can be computed by inserting the corresponding estimate in (13), i.e.,

$$r_j = m_k - \tilde{h}_k \qquad \text{for } \tilde{h}_k \in \tilde{\mathbf{h}}[i]_j; \quad j = \{1, \ldots, p_u\} \tag{18}$$

Next algorithm summarizes the residual evaluation method presented in this paper when sensors in nodes $i, k \in U$ are available.

Algorithm 1 Residual generation

Require:
Matrices: $A_{11}(\mathbf{q})$, N, $\tilde{A}_{12}[i]$, $\tilde{A}_{10}[i]$, A_{21}.
Vectors: $\tilde{\mathbf{h}}_\mathbf{0}[i]$, $\mathbf{d}$.
Measurement of the k-th node pressure: m_k.
Ensure: A set of residuals $\{r_1, \ldots, r_{p_u}\}$.

```
1: for j = 1,...,p_u do
2:     Ã21[j] := A21 without columns j.
3:     d̃[j] := d without element j.
4:     Compute h̃[j] by means of the recursive Newthon-Raphson method in (8)-(9) but using
       Ã10[i], Ã12[i], Ã21[j], h̃0[i] and d̃[j] instead.
5:     h̃k := the k-th element of h̃0[i].
6:     rj := mk − h̃k.
7: end for
```

This algorithm can be executed on-line and used as the residual generator core of the diagnosis system. Besides, since a generic water network model has been considered, the algorithm can be applied to any water distribution network.

5. LEAK LOCALIZATION ANALYSIS

It is important to note that the insensitivity to leak l_j of residual r_j does not mean that the residual must be sensitive to the other leaks. Therefore, a diagnosis analysis is needed in order to determine the expected diagnosis performances from the residuals designed in Section 4.

Because no analytical expression of the residual generator is available, the diagnosis analysis will be based on model properties. Here, with an abuse of notation, the vector of measured pressures is denoted by $\mathbf{m} = \{m_1, \ldots, m_{p_m}\} \subseteq \mathbf{h}$, while the vector of unknown pressures is denoted by $\mathbf{u} = \{u_1, \ldots, u_{p_u - p_m}\} \subseteq \mathbf{h}$. Thus, the on-line information (observations) of the water network available for leak localization consists of the reservoir pressures vector, $\mathbf{h_0}$, the demands vector, $\mathbf{d}$, and the measured pressures vector, $\mathbf{m}$. Then, the set of observations consistent with the leak-free model is defined as

$$\mathcal{O}_\emptyset = \{(\mathbf{h_0}, \mathbf{d}, \mathbf{m}) \mid \exists \mathbf{u}, \mathbf{q} : \text{model } (4) \text{ holds}\} \tag{19}$$

Similarly, the set of observations consistent with a single leak, l_i, model can be defined as

$$\mathcal{O}_{l_i} = \{(\mathbf{h_0}, \mathbf{d}, \mathbf{m}) \mid \exists \mathbf{u}, \mathbf{q}, \mathbf{l} : \text{model } (5) \text{ holds with } \mathbf{l} = (0, \ldots, 0, l_i, 0, \ldots, 0)^T\} \tag{20}$$

Note that this definition can be easy extended to multiple leaks by allowing in (20) that multiple instance of vector $\mathbf{l}$ can take free values.

From the consistency observation sets, leak detectability is next defined.

Definition 1 (Leak detection)
A leak l_i is detectable if $\mathcal{O}_{l_i} \not\subseteq \mathcal{O}_\emptyset$.

Equivalent definitions have been presented for fault detection (see e.g. [12] and [13]). In [14], fault detectability is presented in form of rank condition for linear dynamic systems. Here, we will derive analogous condition but using instead structural matrix properties since the model is non-linear[†]. First, we rewrite the model in (5) as

$$G\mathbf{x} + F\mathbf{l} = H\mathbf{y} \tag{21}$$

where $\mathbf{x} = (\mathbf{q}^T \quad \mathbf{u}^T)^T$, $\mathbf{y} = (\mathbf{h_0}^T \quad \mathbf{m}^T \quad \mathbf{d}^T)^T$ and

$$G = \begin{pmatrix} A_{11}(\mathbf{q}) & A'_{12} \\ A_{21} & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 \\ -I \end{pmatrix},$$

$$H = \begin{pmatrix} -A_{10} & A''_{12} & 0 \\ 0 & 0 & I \end{pmatrix}$$

Matrices A'_{12} and A''_{12} are constructed from A_{12}, by taking into account the columns corresponding to $\mathbf{u}$ and $\mathbf{m}$, respectively.

Now, matrix G is regarded as a structural matrix where all the entries of the diagonal matrix $A_{11}(\mathbf{q})$ are assumed algebraic independent coefficients [15]. Following the idea of [14] and taking into account Definition 1, a leak l_i is detectable if the following condition holds

$$\text{s-rank}([G\ F_j]) > \text{s-rank}(G) \tag{22}$$

where s-rank denotes the structural rank and F_j is the j-th column of matrix F. Note that, according to [15], the structural rank is equivalent to the standard rank under the algebraic independent coefficient assumption.

Leak isolability is analogously defined from the set of consistency observations.

Definition 2 (Leak isolation)
A leak l_j is isolable from a leak l_k if $\mathcal{O}_{l_j} \not\subseteq \mathcal{O}_{l_k}$.

The definition of leak localization comes naturally from Definition 2.

Definition 3 (Leak localization)
A leak l_j can be located if it can be isolated from any other leak l_k , $j \neq k$.

According to [13], condition (22) can be extended to the isolability case. Therefore, a leak l_j is isolable from a leak l_k if

$$\text{s-rank}([G\ F_j\ F_k]) > \text{s-rank}([G\ F_j]) \tag{23}$$

[†]Another possibility, instead of using structural analysis, would be to consider the model linearized at some operating point, then the same results derived here from structural rank conditions will also be valid for standard rank as long as no flow is zero.

Diagnosis analysis can be performed by checking whether conditions (22) and (23) are fulfilled. Note that the analysis does not depend on the residual generator since it is based on model properties. However, it is easy to see that the model (14) used to derive a residual is consistent with this analysis presented in this section.

Proposition 3
The model in (14) can not be used to detect the leak l_j.

Proof
The corresponding G matrix of model (14) is

$$\begin{pmatrix} A_{11}(\mathbf{q}) & \tilde{A}_{12}[i] \\ \tilde{A}_{21}[j] & 0 \end{pmatrix} \tag{24}$$

whereas the corresponding $[G\ F_j]$ matrix of model (14) with leak l_j is

$$\begin{pmatrix} A_{11}(\mathbf{q}) & \tilde{A}_{12}[i] & 0 \\ \tilde{A}_{21}[j] & 0 & 0 \end{pmatrix} \tag{25}$$

From this, it is easy to see that s-rank$([G\ F_i]) =$ s-rank(G) which means that leak l_j cannot be detected. □

This result reconfirm the fact that the decoupling property developed in Section 4 is given by the correct choice of the leak model for each residual computation.

6. CASE STUDY: LIMASSOL WATER NETWORK

The approach proposed in this paper and introduced in Section 4 is applied to a real network (the network of Limassol city in Cyprus) simulated in EPANET as well as compared to the directional residual approach recalled in Section 3. In order to have a reliable comparison, the residuals are averaged out along one day where demands are changed every hour, therefore the residual generator is computed for $L = 24$ as

$$\bar{r} = \frac{1}{L}\sum_{k=1}^{L} r(k) \tag{26}$$

The network of the Limassol city consists of one reservoir, 197 junction nodes ($U = \{1, \ldots, 197\}$) and 239 pipes as shown in Figure 1. Furthermore, it is assumed that three sensors are installed in the following nodes $M = \{75, 81, 152\}$.

6.1. Structured residuals results

The method introduced in Section 4 is here applied to the Limassol water network. First, the localization analysis is performed in order to determine the expected leak localization capability of the method. Then, the residuals are generated from the available sensors in the water network.

6.1.1. Leak localization analysis. The diagnosis analysis introduced in Section 5 is performed on the Limassol network. First, it can be verified that condition in (22) is satisfied for every single leak, which implies that all leaks are detectable. Then, regarding to the isolability analysis, it is verified by applying condition in (23) that 136 of 197 possible leaks can be exactly located. Figure 2 shows the result of the isolability analysis, where a mark in the ij-th position stands for leak in node i can not be isolated from leak in node j. Furthermore, Table I summarizes the information in Figure 2 by showing the leak distribution according its isolability degree. Notice that the worst isolability case concerns four leaks that cannot be isolated among them.

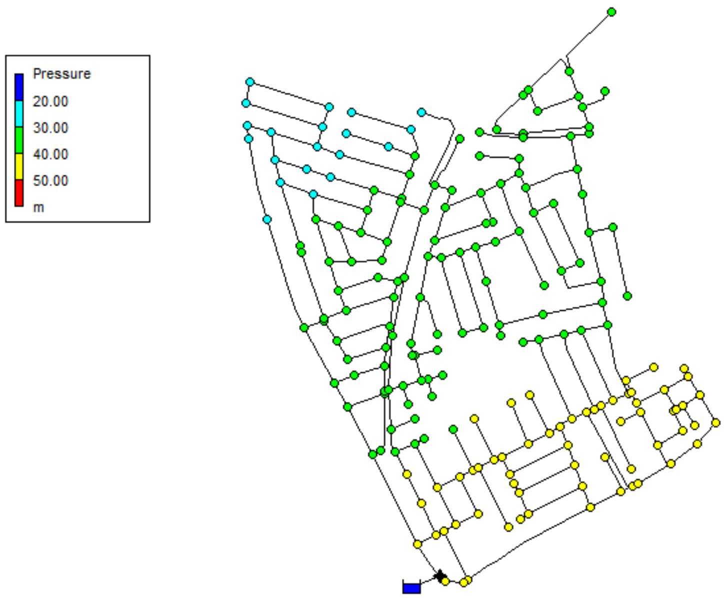

Figure 1. Water network in Limassol, Cyprus

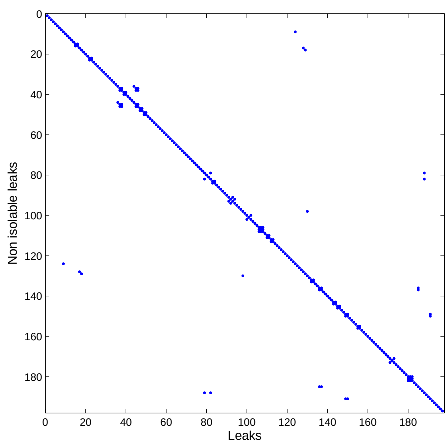

Figure 2. Isolability analysis in Limassol's network

Table I. Number of localizable and partially isolable leaks

	number of leaks
exactly localizable leaks	136
leaks non isolable from one other leak	42
leaks non isolable from two other leaks	15
leaks non isolable from three other leaks	4

6.1.2. Residual generation. Sensor in node 75 is used for leak decoupling according to Proposition 2 with $i = 75$, while sensors in nodes 81 and 152 are used for generating two residual sets, r_i^1 and r_i^2 for $i = \{1, \ldots, 197\}$, according to (13). This means that information provided by two sets of residuals (one from sensor 81 and the other from sensor 152) needs to be evaluated. In the sequel, we will explain how this evaluation is done.

In the case that both residuals r_i^1 and r_i^2 either violate or not the detection threshold at the same time, no further consideration needs to be done. However, the case when only one of this two residuals is exceeding its threshold, while the other is not, needs further analysis. In this later case, the network observations are not consistent with the corresponding l_i-leak model of the violated residual which means that the leak l_i is not feasible. Consequently, a leak in node $i \in U$ is considered present in the network as long as both residuals, r_i^1 and r_i^2, do not violate the detection threshold.

To illustrate this case, we take as example a leak in node 188 which, according to Figure 2, can not be isolated from two other leaks. The results of the residual evaluations are the following:

- Residuals r_i^1 for $i = \{79, 82, 188\}$ do not violate the detection threshold, while the remaining residuals, r_i^1 for $i \neq \{79, 82, 188\}$, do.
- Residuals r_i^2 for $i = \{78, 79, 80, 81, 82, 188\}$ do not violate detection threshold, while the remaining residuals, r_i^2 for $i \neq \{78, 79, 80, 81, 82, 188\}$, do.

From the evaluation procedure detailed above, the diagnosis conclusion to be drawn for this example is that possible leaks are either in node 79, 82 or 188.

6.2. *Directional residuals results*

Directional residuals have also been implemented according to Section 3 for comparison with the approach proposed in this apper. Then, an comprehensive study has been performed by testing the residuals under all possible leak scenarios with different leak magnitudes.

The results presented in Figure 3 show that with the directional residuals it is possible to locate 95.4% of the leaks when the leak magnitude is the same as the one used to compute the sensitivity vectors in (11). However, the performance is deteriorating as long as the leak magnitude is moving far from the one used for sensitivity computations. It can be noticed that the leak localization ratio is reduced near to 50% for a variation of 1lps (liter per second) in the leak magnitude. This demonstrates that directional residuals strongly depends on leak magnitude as it has been mentioned in Section 3.

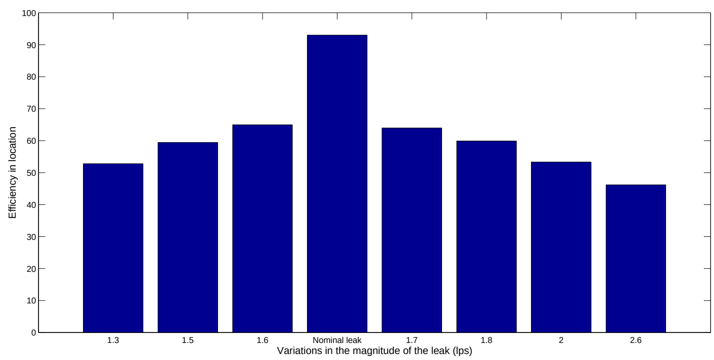

Figure 3. Degradation in localization efficiency using directional residuals with different leak magnitudes.

6.3. *Comparison results between directional and structured residuals*

In the following, we perform a couple of tests in the Limassol's water network with the objective to show how both residual classes perform.

6.3.1. Leak in node 15 with magnitude of 1.67lps. As depicted in Figure 4, the method based on directional residuals is able to find the exact node where the leak takes place, while the method based on structured residuals returns two possible nodes where the leak could take place (one of them is node 15). This is completely consistent with the results obtained in the isolability analysis performed in Section 6.1.1 where leaks in node 15 cannot be exactly located.

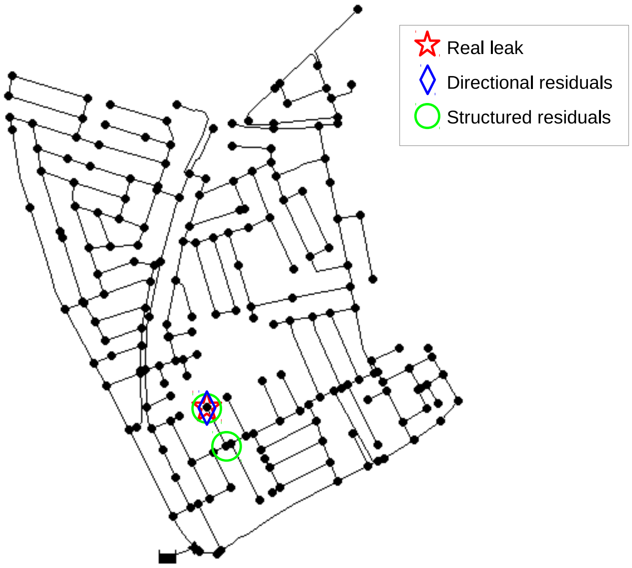

Figure 4. Leak in node 15 with magnitude of 1.67lps

6.3.2. Leak in node 14 with magnitude 2.5lps. In this case, the leak magnitude (2.5lps) is different from the one chosen (1.67lps) to compute the sensitivity vector in (11). Then, the method based on directional residuals fails to locate the leak, whereas the method based on structured residuals locates the right one (see Figure 5). The error of the method based on directional residuals is of one node of distance (the next most plausible leak is the right one). However, the test demonstrates how directional residuals performance depends on the leak magnitude whereas structured residuals do not.

6.4. Directional and structured residuals combination

The results obtained in the previous section suggest that none of both methods is in general better that the other. Therefore, we propose a complementary method in which the localization properties of both methods are combined together. In this new scheme, see Algorithm 2, the results from the structured residuals are first analyzed, and then, in case of more than one possible leak, the directional residuals method is applied in order to know which leak, among all possible leaks, is the most plausible.

Algorithm 2 Directional and structured residuals combination

```
Apply the method based on structured residuals.
if there is more than one possible leak then
    Consider only the possible leaks.
    Apply the method based on directional residuals and
    return the node with the most plausible leak.
else if there is only one possible leak then
    return the node with the possible leak.
end if
```

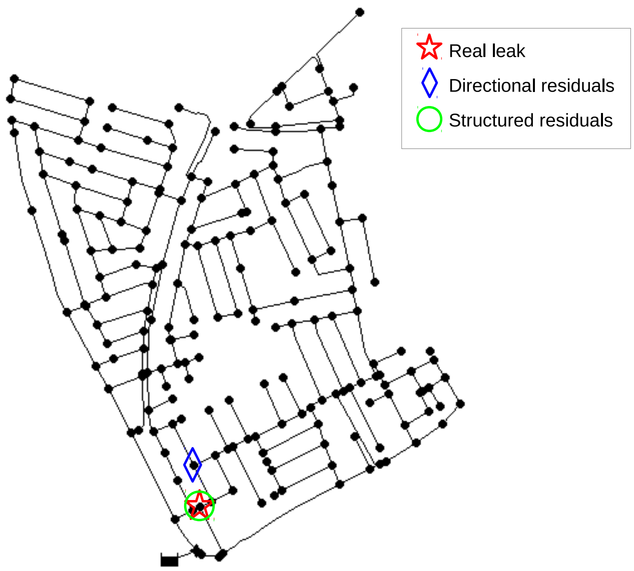

Figure 5. Leak in node 14 with magnitude of 2.5lps

6.4.1. Leak in node 150 with magnitude 3.2lps. This scenario (see Figure 6) clearly shows how leak localization is improved by combining both methods. By applying both methods separately, the method based on structured residuals finds three possible leaks of which one is the right one. On the other hand, the method based on directional residuals returns a wrong leak as the most plausible leak, being the right one at the seven position in the plausibility ranking. However, when both methods are combined the leak is perfectly located. In this case, the most plausible leak from the set of three leaks provided by the structured residuals is the right one.

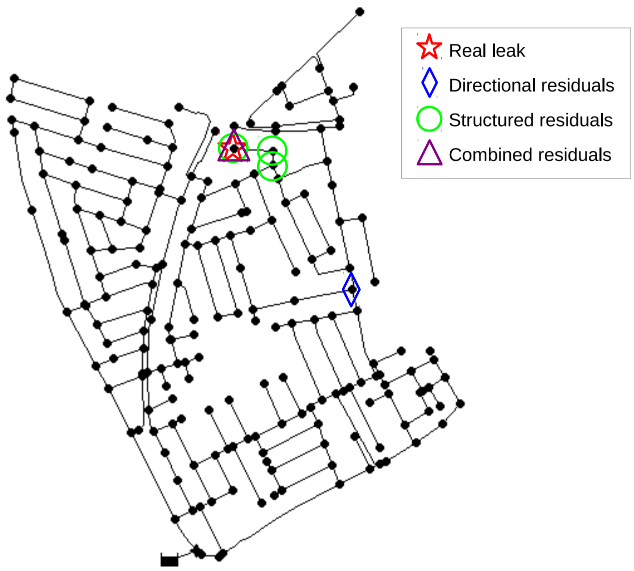

Figure 6. Leak in node 150 with magnitude 3.2lps

6.4.2. Leak in node 36 with magnitude 2.81lps. There are some cases, however, where the combination of the residuals fails in the localization task. For example, in Figure 7, it can be seen how the method based on structural residuals returns two possible leaks and the method based on directional residuals does not find the right one. Consequently, the exact localization of the leak is not achieved even with the combination of both methods.

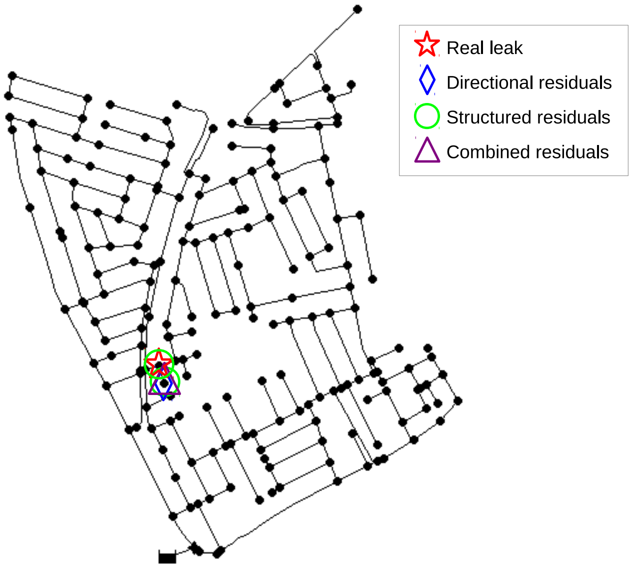

Figure 7. Leak in node 36 with magnitude 2.81lps

However, it is important to note that the best results in the localization task are obtained by using the combination of directional and structured residuals. In fact, structural residuals provide qualitative information by indicating whether a leak is possible. On the other hand, directional residuals provide quantitative information in form of plausibility degree for each possible leak, which a priori is better than the information provided the structural residuals. However, directional residuals suffer of leak magnitude dependability. Therefore, because the advantages of both localization methods are exploited together, no loss of performance can be expected by combining them.

7. CONCLUSIONS

Up to now, because the non-linearities of the model, only sensitivity analysis has been considered so far for model based leak localization in water distribution networks. The present paper goes one step further on the problem of leak localization. Here, a new class of structured residuals, computed numerically, allows to qualitative improve the isolability by decoupling the residuals from specific leaks. However, the authors would like to note that, as it has been shown in the Section 6, the sensitivity analysis is still feasible under this new approach. Thus, the diagnosis designer can take advantage of both approaches at the same time.

Another issue we would like to mention, in favor of the proposed approach, is that the residuals can be computed efficiently. Here, for the sake of space limitation, a small example has been presented. However the numerical tool used to compute the residuals is nowadays widely used to solve water network models involving thousands of nodes. Anyway, in case that computational time was an issue, the approach could be extended in a decentralized fashion were not all residuals wold be computed at once but some of them, and depending on their responses a new set would be selected for computation and so on until the correct leak was located.

The presented work assumes that demands are known and the measurement are exact which is not real in practice. Demands are usually estimated from consumer patterns which are never exact and real pressure sensors are imprecise and with finite resolution. This, together with some parameter inaccuracies, makes that uncertainties should be taken into account for a real implementation. Future research can go in this direction where interval methods and constraint satisfaction tools [16] could be considered.

As mentioned before, in the proposed approach, sensors can be used for two different purposes: leak decoupling and residual generation. This poses a trade-off in the sense that the larger number of sensors for leak decoupling is used, the less residuals are obtained. Moreover, the diagnosis performance depends on the positions and the number of sensors. Therefore, further work should be done in order to known which sensors should be installed and for which purpose.

ACKNOWLEDGEMENT

We would like to express our thanks to Demetrios G. Eliades of the University of Cyprus and to the Water Board of Limassol for helping and allowing to use the case study used in this paper. This work was supported by the Fonds National de la Recherche (FNR) in Luxembourg, under the project CO11/IS/1206050 (SeSaNet), by the Research Chair in Supervision and Advanced Control of Instituto Tecnológico y de Estudios Superiores de Monterrey, and by a CONACYT studentship. It has also been partially grant-funded by CICYT SHERECS DPI-2011-26243, CICYT WATMAN DPI- 2009-13744 of the Spanish Ministry of Education, by iSense grant FP7-ICT-2009-270428 and by EFFINET grant FP7-ICT-2012-318556 of the European Commission.

REFERENCES

1. Lambert A. Accounting for losses: The bursts and background concept. *Water and Environment Journal* April 1994; **8**(2):205–214.
2. Cheung P, Girol GV, Abe N, Propato M. Integrating water systems. chap. Night flow analysis and modeling for leakage estimation in a water distribution system, Taylor and Francis Group: London, UK, 2009; 509–513.
3. Farley T. *Losses in Water Distribution Networks*. IWA Publishing UK: UK, 2003.
4. Colombo A, Lee P, Karney B. A selective literature review of transient-based leak detection methods. *Journal of Hydro-environment Research* April 2009; **2**(2–4):212–227.
5. Pérez R, Puig V, Pascual J, Quevedo J, Landeros E. Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. *Control Engineering Practice* October 2011; **19**(10):1157–1167.
6. Gertler J. *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc.: New York, 1998.
7. Casillas MV, Garza-Castañón L, Puig V. Extended-horizon analysis of pressure sensitivities for leak detection in water distribution networks. *8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes* August 2012; **8**(6):570–575.
8. Pudar L. Leaks in pipe networks. *Journal of Hydraulic Eng.* 1992; **118**(7):1031–1046.
9. Todini E, Pilati S. Computer applications in water supply: vol. 1—systems analysis and simulation. chap. A gradient algorithm for the analysis of pipe networks, Research Studies Press Ltd.: Taunton, UK, UK, 1988; 1–20. URL `http://dl.acm.org/citation.cfm?id=61052.61053`.
10. Collins M, Cooper L, Helgason R, Kennington J, LeBlanc L. Solving the pipe network analysis problem using optimization techniques. *Management Science* 1978; **24**(7):747–760, doi:10.1287/mnsc.24.7.747. URL `http://mansci.journal.informs.org/content/24/7/747.abstract`.
11. Rossman L. Epanet 2 user's manual. *United States Envionmental Protection Agency* 2000; .
12. Blanke M, Kinnaert M, Lunze J, Staroswiecki M. *Diagnosis and Fault-Tolerant Control*. 2nd edn., Springer, 2006.
13. Frisk E, Krysander M, Åslund J. Sensor placement for fault isolation in linear differential-algebraic systems. *Automatica* 2009; **45**(2):364–371.
14. Nyberg M, Frisk E. Residual generation for fault diagnosis of systems described by linear differential-algebraic equations. *Automatic Control, IEEE Transactions on* dec 2006; **51**(12):1995 –2000, doi:10.1109/TAC.2006.884960.
15. Murota K. *Matrices and Matroids for Systems Analysis*. Springer, 2000.
16. L J, Kieffer M, Didrit O, Walter E. *Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics*. Springer-Verlag, 2001.